The Dissertation Committee for Boryana Slavcheva Dimova

certifies that this is the approved version of the following dissertation:

Characterizing Neighborhoods Favorable to Local Search Techniques

Committee:

____

J. Wesley Barnes, Supervisor

____

Elmira Popova, Co-Supervisor

____

John Hasenbein

____

Erhan Kutanoglu

____

Bruce Colletti

Characterizing Neighborhoods Favorable to Local Search Techniques

by

Boryana Slavcheva Dimova, B.S., M.S.

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

December, 2004

Dedication

To Toma and Snezhana

Acknowledgments

I would like to express my sincere appreciation to my advisor, Dr. Wesley Barnes, for his guidance, assistance and support throughout the course of my research. I wish to thank my co-advisor Dr. Elmira Popova for her encouragement and willing support she gave me. Their insight, patience and time in guiding me through my research are very much appreciated.

I give my appreciation to my committee members Dr. John Hasenbein, Dr. Erhan Kutanouglu and Dr. Bruce Colletti.

I owe my deepest gratitude to my son Toma and my daughter Snezhana for keeping those spirits high as we went through this challenging part of our lives. I love them both. Also to my parents and brother for their support and assistance in helping me reach my goal.

I also would like to thank my colleagues and friends for their encouragement through my journey these past three years.

To the reader, I hope that my thesis will assist you with any questions you might come across through your career and maybe will give you challenging ideas.

Characterizing Neighborhoods Favorable to Local Search Techniques

Publication No._____________

Boryana Slavcheva Dimova, Ph.D.

The University of Texas at Austin, 2004

Supervisors: J. Wesley Barnes, Elmira Popova

NP-Complete problems are the most difficult problems to solve and polynomial time algorithms to solve these problems do not exist. One of the more powerful approaches for such problems are heuristic direct search techniques. For a given problem, a landscape is composed of (1) a solution space, (2) an objective function value defined at all elements of the solution space and (3) a direct search neighborhood defined for each element of the solution space.

The goal of the research documented in this dissertation was to extend previous characterizations of landscapes conducive to the success of direct search methodologies. The primary contributions of this dissertation are as follows:

(1) The extension of the characterization of AR(1) elementary landscapes to include arbitrary neighborhood definitions

(2) The creation of an entirely new class of landscapes favorable to direct search methods, a subset of the AR(p) neighborhoods where $p>1$

(3) The development of a lower (upper) bound for a local minima (maxima) in AR(1) elementary landscapes using information stability

(4) The characterization multistep composite AR(1) elementary landscapes

Table of Contents

Notation

NP-hard

COP - Combinatorial Optimization Problem

TSP - Traveling Salesman Problem

$DTMC$ - Discrete Time Markov Chain

X - solution space

f - objective function vector over all X

N - search neighborhood

T - transition matrix associated with $\mathcal{N}$

$\mathcal{L}$ - landscape

L - Laplacian

π – steady state vector associated with T

Π - diagonal matrix with elements $\Pi_{ii} = \pi_i$

α – expectation of f with respect of the transition matrix T

f_{α}. - α-normalized objective function vector

$AR(p)$ - autoregressive process of order p

ε^* - information stability

d_{ii} - degree of node i

D - degree matrix of A

μ – arithmetic average of all f_i in X

∇^2 - Grover's Difference Operator

Δ - Stadler's matrix operator

$\underset{y \in N(x)}{Avg} f(y)$ - average objective function value of its neighbors

STSP - Symmetric Traveling Salesman Problem

a_t - uncorrelated random variable with mean zero and constant variance

$E[.]$ - expectation

$\mathrm{var}[.]$ - variance

$\rho_\pi(\mathrm{s})$ - autocorrelation function at lag s

F - probability of first passage time

λ - eigenvalue of L

N_A - one-step neighborhood with adjacency matrix A

N_{A^2} - two-step neighborhood with adjacency matrix A^2

G - square matrix where $G_{ij}=1$

L_2 - Laplacian associated with T^2

Chapter 1

Introduction

The research documented in this dissertation is directed at analyzing and investigating the properties of direct search neighborhoods used in heuristic and metaheuristic methods for the solution of combinatorial optimization problems.

1.1 Elementary Landscapes and Their Properties

A *landscape* for a combinatorial optimization problem (COP) is defined by $\mathcal{L} = (X, f, N)$, where X is the solution space, f is the objective function vector over all X, N is the search neighborhood and T is a *transition matrix* associated with N. The Laplacian is defining as $L=I-T$. The vector $\pi = [\pi_i]$ is the steady state vector associated with T and $\alpha = \pi' f$ is the expectation of f with respect to the transition matrix T. The *α-normalized* objective function vector is defined as $f_\alpha = [f-\alpha]$.

Grover (1992) defined a *wave equation* using a difference operator, which is a discrete analog of a continuous Laplacian operator. Stadler (1996) defined a Laplacian operator and proved, in the case of regular and symmetric neighborhoods, the equivalence of the two operators. Landscapes that satisfy the wave equation are *elementary landscapes* which have favorable properties for direct search approaches to COPs.

In this dissertation, a *general* Laplacian operator is defined and shown to be equivalent to the operators proposed by Grover (1992) and Stadler (1996) for arbitrary search neighborhoods. Several properties of composite neighborhoods constructed upon elementary landscape are also presented.

1.2 Random Walks on Landscapes

A random walk on a landscape is defined as a stochastic process over the solution space where we randomly start at some solution and move to a neighbor defined by *N*. The objective function values of the visited solutions yield a univariate time series which can be studied to capture statistical properties of the landscape.

Weinberger (1990) described a *random walk* on a landscape. The simplest correlated case is AR(1) process. Stadler et al. (1993) proved that when an elementary landscape has a regular and symmetric neighborhood, a random walk on such a landscape will be consistent with an AR(1) process, i.e., the associated autocorrelation function will be governed by an exponential function. This dissertation generalizes this result to embrace arbitrary elementary landscapes and develops a Characteristic Landscape Equation for AR(p) processes, where $p \geq 2$.

1.3 Lower (upper) Bounds for Local Minima (Maxima) of Elementary Landscape Using Information Theory

Information theory is a different approach to analyzing the properties of landscapes using two entropy measures of the associated time series, partial information and information stability. This method views landscapes as an ensemble of objects that are related to the neighborhood structure. The information content of a set of objects is a measure of how difficult it is to describe the set, i.e., it defines a metric on the ruggedness of the set. In this research a lower (upper) bound for a local minima (maxima) of elementary landscape is developed using information stability, ε^*, the eigenvalue of the Laplacian λ and the expectation of f with respect of the transition matrix T.

Chapter 2

Literature Review

2.1 Heuristics

Many NP-hard planning and management problems, embedded in practical arenas such as logistics, communications, and manufacturing, have motivated the development of advanced solution methodologies. Classical optimization algorithms do not work well for such problems due to the exponential size of their solution spaces. Unlike optimal algorithms, local search metaheuristic techniques, like tabu search, do not guarantee optimal solutions. Empirical results have shown that, in many cases, such heuristics find high quality solutions with dramatically less computational effort relative to optimal algorithms.

In this chapter, the relevant previous published research is reviewed and mathematical foundations are presented sufficient for the reading of the remaining chapters of this dissertation.

2.2 Local Search Algorithms

A local search algorithm is defined when rules for choosing an incumbent solution, a search neighborhood, a decision rule for new solution selection and termination criteria are stipulated. A local search algorithm starts from an initial incumbent solution and moves to a neighboring solution based on the selection rule.

In the context of a local search procedure applied to a real combinatorial optimization problem (COP), a move is an operation on an incumbent solution $x \in X$ (the solution space) that transforms x into a neighboring solution, y. A stipulated set of such moves for each and every $x \in X$ defines the search neighborhood.

The local search procedure repeats until a termination criterion is satisfied. For a specified problem, some neighborhoods work very well and a good solution is found relatively quickly. Unfortunately, other neighborhoods may be poorly suited to the problem and yield disappointing results. In the early 1990's, research directed at the study of search neighborhoods and their associated problem landscapes began (Grover (1992), Codenotti and Margara (1992)).

2.3 The Definition of a Landscape

A landscape is defined by the triplet $L = (X, f, \mathcal{N})$ (Barnes et al. (2003)), where $X = [x_i]$ is the solution space, $f = [f(x_i)] = [f_i]$ is the objective function vector over all X, and N is the search neighborhood defined by a digraph where the nodes represent the $x_i \in X$. The neighborhood digraph has an associated adjacency matrix A and transition matrix T. For each $x_i \in X$, a nonzero a_{ij} (a positive integer) designates x_j as a neighbor of x_i. The transition matrix is defined

to be $T = [\,t_{ij}\,] = [a_{ij}/d_{ii}]$ where $d_{ii} = \sum_{\forall j} a_{ij}$, the degree of node i. The matrix $D = [d_{ij}]$ (where $d_{ii} = \sum_{\forall j} a_{ij}$) the degree matrix of A. The δ-normalized objective function vector is defined to be $f_\delta = [\, f(x_i) - \delta \,] = [\, f_{\delta i} \,]$. The arithmetic average of all f_i in X is denoted as μ.

2.4 Grover's Difference Equation

Grover (1992) defined the following difference equation

$$\nabla^2 f(x) + \frac{k}{n} f(x) = 0, \forall x \in X \tag{2.1}$$

with constant $k > 0$ and noted its similarity to the "wave equation" of mathematical physics. ∇^2 is the difference operator

$$\nabla^2 f(x) = \sum_{y \in N(x)} \frac{(f(y) - f(x))}{d(x)} \tag{2.2}$$

which is the neighborhood average of objective function value deviations, i.e., the $f(y) - f(x)$, where x is an incumbent solution, the $y \in \mathcal{N}(x)$ are the one step neighbors of x and $d(x)$ is the cardinality of the neighborhood of x.

2.5 The Laplacian Equation

Considering *only* regular and symmetric N, i.e., neighborhoods with symmetric A with common degree d for all $x_i \in X$, Stadler (1996) develops a matrix version of Grover's difference equation by defining the matrix operator

$$\Delta = A - dI$$

where I is an $|X| \times |X|$ identity matrix. For regular and symmetric neighborhoods $\Delta \equiv d\nabla^2$. Thus Equations 2.1 and 2.3,

$$\Delta f + \lambda f = 0 \tag{2.3}$$

are equivalent.

For landscapes with regular and symmetric neighborhoods, a complete orthogonal set of eigenvectors of Δ exists (Stadler (1996), Angel and Zissimopoulos (1999)). This allows any objective function vector f to be expanded as a Fourier series in terms of eigenvectors of Δ, a property essential to many of Stadler's proofs (Stadler (1996)).

2.6 Elementary Landscapes

An *elementary* landscape is a landscape that satisfies Equations 2.1 and 2.3 (Stadler (1996)). Grover (1992) studied landscapes that satisfy Equation 2.1 where $0 \le \frac{k}{n} \le 1$. While Grover's definition allows landscapes with arbitrary neighborhood graphs, all of his results require regular and symmetric neighborhood graphs. Grover (1992) showed that elementary landscapes have two properties favorable to local search:

(1) In an elementary landscape, local optima are superior to the average value μ of the objective function over the solution space

(2) The number of steps to reach a solution at least as good as μ from any starting point is linear in the problem size.

Grover (1992) also shows that landscapes arising from certain well-known COPs (the symmetric traveling salesman problem (TSP), the min–cut graph partitioning problem, the graph coloring problem, the not-all-equal satisfiability problem and weighted partition problem) can satisfy Equation 2.1 when "proper" neighborhoods are selected.

Codenotti and Margara (1992) extend Grover's (1992) results for the symmetric TSP, developing new neighborhoods that satisfy Equation 2.1 and deriving additional structural properties for elementary landscapes. For any solution, x, define $\underset{y \in N(x)}{Avg}\ f(y)$ to be the average objective function value of its neighbors, and μ to be the average objective function value over the solution space. One and only one of the following relations is true:

$$\begin{aligned} f(x) &< \underset{y \in N(x)}{Avg}\ f(y) < \mu \\ f(x) &= \underset{y \in N(x)}{Avg}\ f(y) = \mu \\ f(x) &> \underset{y \in N(x)}{Avg}\ f(y) > \mu \end{aligned} \tag{2.4}$$

Thus, $\underset{y \in N(x)}{Avg}\ f(y)$ is bounded between $f(x)$ and μ. This implies that "on average" all $x \in X$ have neighbors whose *f(y)* are "similar" to *f(x)* (on the same side of μ). Like Codenotti and Margara (1992), Stadler (1996) also observes this relation and characterized such landscapes by smooth "rolling hills and valleys". As a consequence of their assumption of regular and symmetric neighborhoods, Grover (1992), Codenotti and Margara (1992), Stadler (1993,1996, 2000) force the normalization parameter to be μ.

Colletti (1999) proved that symmetric multiple TSPs with neighborhoods defined on an arbitrary collection of exchange neighborhoods also satisfy Equation 2.1. Barnes and Colletti (2001) investigate a large number of neighborhoods that satisfy Grover's difference equation for all symmetric traveling salesman problems (STSPs). Solomon et al. (2003) extend the previous work on STSPs to a much larger class of TSPs with weakly symmetric distance matrices.

2.7 Random Walks on Landscapes

Previous research has shown the efficacy of imposing an elementary landscape upon the search topology. However, the exponentially large solution spaces of practical COPs *require* the use of statistical sampling methods to analyze such landscapes.

Weinberger (1990) propose a *random walk* on the landscape as a statistical method to analyze the landscapes. Given a randomly selected initial incumbent solution, the walk randomly moves to a neighboring solution with probability t_{ij}. The new solution becomes the incumbent solution and the process continues until a termination criterion is satisfied. An important assumption for Weinberger's model is that the landscape is statistically isotropic (Stadler and Happel (1999)), i.e., the statistical properties of the random walk are independent of the starting point and the f_i of this sequence form a stationary random process for the assumed joint distribution of *f*. Weinberger (1990) characterizes a landscape using the *autocorrelation function* and presumes that the random walk steps from solution *i* to solution *j* according to an ergodic Markov chain with probability transition matrix *T*. Weinberger (1990), Weinberger and Stadler (1993), and Stadler (1996) consider only regular and symmetric *T* which forces that the steady state vector, π, of the associated Markov chain to be uniform.

Hordijk and Stadler (1998) use a random walk to calculate the correlation function *f* to assess the "ruggedness" of the landscape. Stadler (1995, 1996) partitions the assumed distance transitive neighborhood graph to define a compressed graph of much smaller size based on that partitioning. A random walk on the new graph is then proposed for the landscape analysis. Hordijk (1996) uses Weinberger's random walk and Box-Jenkins time series analysis to measure and express the correlation structure of landscapes. Vassilev, et al. (2000) use information theory to analyze the time series generated from a random walk.

2.8 Autoregressive Prosesses and Landscapes

A General Linear Process is defined as (Box and Jenkins (1976))

$$\tilde{z}_t = \varphi_1 \tilde{z}_{t-1} + \varphi_2 \tilde{z}_{t-2} + \ldots + a_t \tag{2.5}$$

where $\tilde{z}_t = z_t - \mu$. The a_t are uncorrelated random variables with mean zero and constant variance, i.e. $E[a_t] = 0, \mathrm{var}[a_t] = \sigma^2$. Consider the special case of Equation 2.5 when only the first *p* coefficients are nonzero,

$$\tilde{z}_t = \varphi_1 \tilde{z}_{t-1} + \varphi_2 \tilde{z}_{t-2} + \ldots + \varphi_p \tilde{z}_{t-p} + a_t \tag{2.6}$$

The process defined in Equation 2.6 is an *autoregressive process* of order *p* (AR(p)) with autocorrelation function

$$\rho_k = \varphi_1 \rho_{k-1} + \varphi_2 \rho_{k-2} + \ldots + \varphi_p \rho_{k-p}\text{, for } k>0 \tag{2.7}$$

If $p=1$, we have an AR(1) or *Markov* process, i.e., the state at time *t* depends only of the state at time *(t-1)*.

Weinberger (1990) studied how the autocorrelation function of *f* obtained from a random walk on regular and symmetric neighborhoods can characterize the landscape. Weinberger (1990) presents a wide class of landscapes, including the *N-k* model and the TSP, where the time series associated with a random walk is consistent with an AR(1) process, i.e., the associated autocorrelation function governed by an exponentially declining function. Weinberger denoted them as AR(1) landscapes.

Considering only regular and symmetric *T*, Stadler, Seitz and Wagner (2000) show that $\mathcal{L}$ is elementary if and *only* if the autocorrelation function is exponentially declining, i.e., consistent with an AR(1) landscape.

2.9 Information Theory and Landscapes

Information theory is a branch of mathematics dealing with the efficient and accurate storage, transmission, and representation of information. Mathematically, the amount of disorder of the system can be calculated using the Shannon definition of entropy of a variable x

$$H(x) = -\sum_{x} P(x) \log_2 [P(x)] \tag{2.8}$$

where $P(x)$ is the probability that X is in the state x and $P \log_2 P$ is defined as 0 if $P = 0$.

Vassilev, et al. (2000) propose an information analysis of landscapes. This idea is inspired by the concept that the information content of an individual system is a measure of how difficult it is to describe that system. They consider a landscape as an ensemble of objects that are related to the neighboring points. They propose a random walk on the landscape and define information content, partial information content and information stability as information characteristics of the ensemble. Vassilev, et al. (2000) analyze the structure of selected landscapes using these parameters. A variety of "shapes" on the landscape relate to the local neighborhood about a landscape point and they can be captured by the information content. In particular, the upper bound of the magnitude of the landscape optima can be obtained using information stability.

We will give a brief description of this process. The constant ε is a nonnegative real number. Let $\{f_t\}_{t=0}^{n}$ be the time series generated from a random walk (Weinberger 1990). First transform the time series using the mapping

$$S(\varepsilon) = \{s_1, s_2, \ldots, s_n\} \tag{2.9}$$

where

$$s_i = \begin{cases} \bar{1} & \text{if} \quad f_\alpha(x_i) - f_\alpha(x_{i-1}) \le -\varepsilon \\ 0, & \text{if} \quad |f_\alpha(x_i) - f_\alpha(x_{i-1})| \le \varepsilon \\ 1, & \text{if} \quad f_\alpha(x_i) - f_\alpha(x_{i-1}) \ge \varepsilon \end{cases}$$

The idea behind this transformation is to be able to extract information from the landscape by ignoring some non-essential features. The value of ε measures the accuracy of the calculations of the string (2.9).

Vassilev et al. (2000) characterize the ruggedness or "information content" of the landscape by introducing an entropy measure of the ensemble associated with sub-blocks of length two of the string (2.9). In addition, they measure the ruggedness of the landscape by the modality of the time series path using the following construction. Consider a compression of S(ε) deleting all 0 values and all the elements whose right adjacent element is of equal value. This yields a new set which has the form $\{\bar{1}, 1, \bar{1}, \ldots,\}$ and is the shortest string that represents the slopes of the neighboring landscape path. The length of the compressed string is the modality, *M.* The "partial information content" is defined as $M(\varepsilon) = \frac{M}{n}$, $0 \le M(\varepsilon) \le 1$, where $M(\varepsilon) = 0$ implies a flat landscape and $M(\varepsilon) = 1$ implies maximal modality. The relative accuracy of the estimation of the information content and partial information content is inversely proportional to ε. *Information stability* is characterized by the smallest value of ε, ε^*, such that $S(\varepsilon)$ is a string of zeros.

This completes the mathematical foundations and literature review. The following chapters present the new findings obtained during this dissertation research.

Chapter 3 presents the proof of uniqueness of the normalization constant for elementary landscapes, equivalency between Grover's difference equation and extended Laplacian equation, smoothing rugged landscapes and provide some additional properties of elementary landscapes such as: bounds on local optima, and the normalization constant for elementary landscapes with doubly stochastic transition matrix.

Chapter 4 presents the proof of equivalency between an elementary landscape and a univariate time series on f_α generated from a random walk on T, consistent with an AR(1) process in the case of a general neighborhood and four new properties of objective function vectors associated with elementary landscapes. Chapter 5 discusses the composition of elementary landscapes and proves that such landscapes are also elementary. The characteristics of composite elementary landscapes (smooth or rugged) is then considered. Chapter 5 concludes with the development of the first non trivial lower(upper) bound of a local minimum(maximum) for the elementary landscapes.

Chapter 6 develops the Characteristic Landscape Equation for an AR(p) landscape, gives a proof of the equivalency between a landscape that satisfies the corresponding Characteristic Landscape Equation and a univariate time series on f_α generated from a random walk on T, consistent with an AR(p) process. Chapter 6 also investigates the properties of AR(2) landscapes and identifies, under additional parametric conditions, a set of landscapes which have properties favorable for local search. These are the first new landscapes, in addition to Laplacian elementary landscapes that have been shown to be favorable neighborhood for a local search.

Chapter 7 gives directions for further research.

Chapter 3

The Theory of Elementary Landscapes

This chapter proves and discusses the following contributions:

(1) Grover's difference equation and an *extended* Laplacian equation are equivalent.

(2) The objective function normalization constant for an elementary landscape is unique.

(3) Two types of elementary landscapes, smooth and rugged, exist.

(4) Rugged elementary landscapes yield smooth elementary landscapes under a two move composite neighborhood.

These contributions are accompanied by the proofs of some additional properties such as bounds for local optima, the definition of smooth and rugged elementary landscapes, the analysis of extreme eigenvalues for elementary landscapes, and the derivation of the normalization constant for elementary landscapes with doubly stochastic transition matrices. While combinatorial optimization is a focus of this chapter, the definition of a landscape presented here is completely abstract and the results in this chapter apply to the many situations in which landscapes occur.

3.1 Equivalence Between Grover's Difference Equation and the Extended Laplacian Equation

Using a graph Laplacian similar to that proposed by Chung (1997), we extend Stadler's definition of elementary landscapes to arbitrary neighborhood digraphs. This allows us to characterize a more general class of landscapes satisfying Grover's wave equation. After existing results on elementary landscapes are extended to this larger class, two general types of elementary landscapes are discussed.

For any digraph associated with an adjacency matrix A we define the *extended* Laplacian by

$$L = I - D^{-1}A, \tag{3.1}$$

where $D^{-1}A$ is a stochastic matrix, i.e., all row sums are equal to 1, and so its possibly complex eigenvalues have modulus in the interval [-1, 1] (Stewart(1994)). Consequently, the eigenvalues of L have modulus in the interval [0, 2].

Lemma 3.1: For any $f: X \to R$, $D^{-1}Af(x)$ is the average value, $\underset{y \in N(x)}{Avg}\ f(y)$, of f on $N(x)$, i.e., $D^{-1}Af(x) = \underset{y \in N(x)}{Avg}\ f(y)$.

Proof: Routine. □

Definition 3.1: An elementary landscape is one in which $f_\alpha = [f_i - \alpha]$ is an eigenvector of L for some real number α, i.e.

$$Lf_\alpha = \lambda f_\alpha \tag{3.2}$$

Theorem 3.1: The linear operators L on the space of functions *f: X* $\rightarrow$ *R* are equivalent to $(-\nabla^2)$, i.e. $L \equiv -\nabla^2$.

Proof: The proof is a direct application of the result from Lemma 3.1**.** $\square$

The following corollary provides the promised characterization:

Corollary 3.1: A landscape *(X, f, N)* is elementary if and only if (X, f_α, N) satisfies Grover's wave equation (2.1) for some real α.

With this characterization, the classical mathematical tool of spectral analysis can now be applied in studying different types of landscapes for a given neighborhood.

3.2 Uniqueness of the Normalization Constant for Elementary Landscapes

We now show that if the normalization constant, α, satisfying Equation 3.2 exists, it is unique.

Theorem 3.2: For a fixed *non-flat* landscape $\mathcal{L}$ = *(X, f, N)* if there exists a normalization constant α such that the landscape $\mathcal{L} = (X, f_\alpha, N)$ is elementary then α is unique.

Proof: Define $e = (1, 1, ..., 1)'$ and $0 = (0, 0, ..., 0)'$ to be $|X|$ dimensional vectors of ones, and zeros, respectively. Let $\alpha \neq \beta$ be two different normalization constants yielding elementary landscapes, i.e.,

$$Lf_\alpha = \lambda_\alpha f_\alpha \tag{3.3}$$

and

$$Lf_\beta = \lambda_\beta f_\beta \tag{3.4}$$

Subtracting (3.4) from (3.3) we obtain

$$\lambda_\alpha f_\alpha - \lambda_\beta f_\beta = L(f_\alpha - f_\beta) = (\beta - \alpha)Le = 0$$

Therefore, $\lambda_\alpha f_\alpha = \lambda_\beta f_\beta$.

We must consider two possibilities: $\lambda_\alpha = \lambda_\beta$ and $\lambda_\alpha \neq \lambda_\beta$.

- Let $\lambda_\alpha = \lambda_\beta = \lambda$ where $\lambda \neq 0$ (the proposition disallows flat landscapes). Hence $f_\alpha = f_\beta \Rightarrow \alpha e = \beta e \Rightarrow \alpha = \beta$.
- Let $\lambda_\alpha \neq \lambda_\beta$ ($\lambda_\alpha \neq 0, \lambda_\beta \neq 0$)

  Hence $f_\alpha = \dfrac{\lambda_\beta}{\lambda_\alpha} f_\beta \Rightarrow f - \alpha e = \dfrac{\lambda_\beta}{\lambda_\alpha}(f - \beta e)$

$$\Rightarrow (1 - \frac{\lambda_\beta}{\lambda_\alpha})f = (\alpha - \frac{\lambda_\beta}{\lambda_\alpha}\beta)e$$

$$\Rightarrow f = [(\alpha - \frac{\lambda_\beta}{\lambda_\alpha}\beta)\frac{\lambda_\alpha}{\lambda_\alpha - \lambda_\beta}]e$$

which implies that all f_i are equal, yielding a flat landscape, a contradiction of the assumption in the proposition. Therefore if the normalization constant for a fixed f exists, it is unique. □

Hence, given a fixed λ, previously studied neighborhoods which are known to have elementary landscapes for f_μ can not yield elementary landscapes for f_α where $\alpha \neq \mu$ (Weinberger (1990), Grover (1992), Stadler (1996)).

3.3 Two Types of Elementary Landscapes

This section first explains the properties of elementary landscapes, which make them amenable to local search, and then characterizes two classes of elementary landscapes, resulting from varying the parameter λ: smooth and rugged elementary landscapes.

While it is possible to define COPs where *f* can achieve complex values, for the purposes of this work we limit ourselves to COPs (and therefore landscapes) where *f* can achieve only real values. This limits us to elementary landscapes where both the eigenvector *f* and therefore the eigenvalue λ are real, since *L* is real by definition. As discussed in Chung (1997), the eigenvalues of *L* lie in the interval [0, 2], so that an elementary landscape can only exist for $\lambda \in [0, 2]$.

3.3.1 Upper (Lower) Bounds for Local Minima (Maxima) in Elementary $\mathcal{L}$

A solution *x* is a local minimum if *f(x)* ≤ *f(y)* for every $y \in N(x)$. The notion of local maximum is defined similarly.

Theorem 3.3: In an elementary landscape with $\lambda > 0$, local minima have values at most α, and local maxima have values at least α.

Proof: Equation 3.2 with Lemma 3.1 yields, for each $x \in X$,

$$\underset{y \in N(x)}{Avg}\ f_\alpha(y) = (1-\lambda) f_\alpha(x) \tag{3.5}$$

If x^* is a local minimum, then its objective value doesn't exceed those of its neighbors, implying

$$f_\alpha(x^*) \leq \underset{y \in N(x^*)}{Avg}\ f_\alpha(y) = (1-\lambda) f_\alpha(x^*) \tag{3.6}$$

From this we immediately have $\lambda f_\alpha(x^*) \leq 0$, and since λ>0, $f(x^*) \leq \alpha$ as required. The corresponding proof for local maxima is symmetrical. □

In the next section we discuss the properties of elementary $\mathcal{L}$ with $\lambda = 0$ and $\lambda = 2$.

3.3.2 Extreme Eigenvalues for Elementary Landscapes

If for each *x, y* $\in$ *X* there is a directed path from *x* to *y,* the directed multigraph that the neighborhood *N* defines is *connected.* The following proposition explains why elementary landscapes with eigenvalue *0* or *2* are often degenerate and permits us to restrict further investigations to elementary landscapes with $\lambda \in (0, 2)$.

Theorem 3.4: If *N* defines a connected digraph, then:

(i) An elementary landscape with $\lambda = 0$ has constant objective function, i.e. a flat landscape where *f(x)* has the same value $\forall x \in X$;

(ii) If in addition $x \in N(x)$ for at least one $x \in X$, then there is no elementary landscape with $\lambda = 2$.

Proof: First notice that corresponding to an eigenvalue λ of L is an eigenvalue $1-\lambda$ of $D^{-1}A$. If the neighborhood digraph is connected then $D^{-1}A$ is a nonnegative irreducible matrix (Kulkarni (1995)).

(i) Suppose f_α is an eigenvector for the eigenvalue $\lambda = 0$ of L. Then the corresponding eigenvalue of $D^{-1}A$ is 1, which is the eigenvalue of maximum modulus. By the Perron-Frobenius Theorem, 1 is a simple root of the characteristic equation of $D^{-1}A$ which implies that it is associated with a single eigenvector. Since the vector whose entries are all equal to 1 is an eigenvector of L with eigenvalue $\lambda = 0$, we conclude that this is the only eigenvector corresponding to that eigenvalue. Hence, the objective function $f = [1 + \alpha]$ is a constant vector.

(ii) If in addition $x \in N(x)$ for some $x \in X$ then the irreducible $D^{-1}A$ has a positive trace, which implies that $D^{-1}A$ is a primitive matrix. Once again, by the Perron-Frobenius Theorem (Cvetkovic (1980)), such a matrix has unique eigenvalue of maximum modulus. Therefore $1-\lambda \neq -1$, i.e., $\lambda \neq 2$. □

Next, we identify two types of elementary landscapes depending on the value of $\lambda \in (0, 2)$.

3.3.3 Smooth Elementary Landscapes

If $0 < \lambda \leq 1$, then Equation 3.5 implies

$$f_\alpha(x) \leq \underset{y \in N(x)}{Avg}\, f_\alpha(y) \leq 0 \text{ (i.e., } f(x) \leq \underset{y \in N(x)}{Avg}\, f(y) \leq \alpha\text{)}, \qquad \text{for } f_\alpha(x) \leq 0$$

and

$$f_\alpha(x) \geq \underset{y \in N(x)}{Avg}\, f_\alpha(y) \geq 0 \text{ (i.e., } f(x) \geq \underset{y \in N(x)}{Avg}\, f(y) \geq \alpha \text{)}, \qquad \text{for } f_\alpha(x) \geq 0$$

This implies that, on average, all $x \in X$ have neighbors whose *f(y)* are similar to *f(x)* (on the same side of α), i.e., the landscape is characterized by smooth "rolling hills and valleys" as we discussed in Section 2.6. However, in the general case where α need not equal μ, stronger information may be inferred. For example, if $\alpha < \mu$, then local minima have lesser values than when $\alpha = \mu$.

3.3.4 Rugged Elementary Landscapes

If $1 \leq \lambda < 2$, then Equation 3.5 implies that

$$f_\alpha(x) \leq 0 \leq \underset{y \in N(x)}{Avg}\, f_\alpha(y), \qquad \text{for } f_\alpha(x) \leq 0$$

and

$$f_\alpha(x) \geq 0 \geq \underset{y \in N(x)}{Avg}\, f_\alpha(y) \geq 0, \qquad \text{for } f_\alpha(x) \geq 0$$

This relation implies that all solutions are surrounded, on average, by solutions with values on the opposite side of α.

At the intersection of these two cases (for $\lambda = 1$), we have

$$\underset{y \in N(x)}{Avg}\, f_\alpha(y) = 0 .$$

While rugged-elementary landscapes would present a more difficult challenge for a simple greedy local search, knowledge that such a landscape is present would enhance the strategic search possibilities of more sophisticated reactive and adaptive metaheuristic approaches like tabu search. The identification of rugged-elementary landscapes extends the incomplete results of Codenotti and Margara

(1992) and Stadler (1996) who apparently were only aware of smooth-elementary landscapes.

3.4 Smoothing Rugged Landscapes

In this section we limit our attention to regular adjacency matrices and show that rugged elementary landscapes can be smoothed by using the corresponding two-step neighborhood. Empirical evidence indicates that Glover and Laguna's (1997) ejection chain method of generating compound neighborhoods enhances a metaheuristic search method. The findings in this study may provide theoretical insight into the success of such techniques. The investigation of neighborhoods that always yield smooth landscapes is an important area of research.

Let us assume that the adjacency matrix A defines a regular neighborhood N_A, which in turn yields an elementary landscape. We now consider cases where the associated two-step neighborhood N_{A^2} must yield a smooth ($\lambda < 1$) landscape.

Theorem 3.5: Every regular neighborhood N_A defines a two-step regular neighborhood N_{A^2} such that if *(X, f,* N_A*)* is an elementary landscape with a real eigenvalue, then (X, f, N_{A^2}) is a smooth elementary landscape.

Proof: Let d be the degree of regularity of the neighborhood N_A. It is easy to see that N_{A^2} is regular with degree d^2. The Laplacian L_A of N_A is $L_A = I - (1/d)A$, while the Laplacian of N_{A^2} is $L_{A^2} = I - (1/d^2)A^2$.

Suppose *(X, f, N_A)* is elementary. Then

$$\lambda f_\alpha = L_A f_\alpha = (I - \frac{1}{d} A) f_\alpha$$

so that

$$Af_\alpha = d(1-\lambda) f_\alpha .$$

Therefore

$$L_{A^2} f_\alpha = (I - \frac{1}{d^2} A^2) f_\alpha = f_\alpha - \frac{d^2}{d^2}(1-\lambda)^2 f_\alpha = (1-(1-\lambda)^2) f_\alpha$$
$$= (2\lambda - \lambda^2) f_\alpha$$

so that *(X, f, N_{A^2})* is elementary with eigenvalue *($2\lambda - \lambda^2$)*. Since $2\lambda - \lambda^2$ has maximum value *1*, *(X, f, N_{A^2})* is smooth. □

Recalling that every real symmetric matrix has real eigenvalues we have the following:

Corollary 3.2: If *A* defines a regular, symmetric neighborhood, then for every elementary landscape *(X, f, N_A)* the two-step landscape *(X, f, N_{A^2})* is smooth.

If the one-step adjacency matrix, *A*, is irregular and symmetric then a smooth two-step landscape exists. To see this, note that A^2 is positive semidefinite. Hence, $D_2^{-1}A^2$ is also positive semidefinite as discussed in Lancaster and Tismenetsky (1985), where D_2 is the two-step diagonal degree matrix. Therefore, the eigenvalues, λ_2, of the two-step Laplacian $L_{A^2} = I - D_2^{-1}A^2$ correspond to the *($1 - \lambda_2$)* eigenvalues of $D_2^{-1}A^2$, which are nonnegative, hence $\lambda_2 \leq$ *1*. Unfortunately, unlike the case where *A* is regular, f_α is not necessarily invariant when one moves from the one-step to the two-step neighborhood. We investigate this question in further detail in Chapter 5.

3.5 Doubly Stochastic Transition Matrices and the Normalization Constant

In Section 2.6 we presented the fact that previous researchers were not aware that the normalization constant α could be different from μ. In this section we discuss this fact and present an example for elementary landscape where $\alpha \neq \mu$.

Theorem 3.6: If the adjacency matrix is regular and symmetric then the eigenvalues and eigenvectors of the Laplacian must be real.

Proof: As noted earlier, the Laplacian is clearly symmetric if the adjacency matrix is. The result now follows immediately from the well-known fact that the eigenvalues and eigenvectors of a symmetric matrix are real. □

In fact, it can easily be shown that $\alpha = \mu$ for a larger class of landscapes which includes those with regular and symmetric neighborhood graphs as a subclass. That larger class is the set of those whose normalized adjacency matrices $D^{-1}A$ are doubly stochastic, i.e. the row and column sums are all unity.

Theorem 3.7: If *(X, f, N_A)* is (α, λ)-elementary with $D^{-1}A$ doubly stochastic, then $\alpha = \mu$, the global average of *f*.

Proof: Recall, a matrix is doubly stochastic if its rows and columns all sum to *1*. Therefore, if $D^{-1}A$ is doubly stochastic then the columns of $L = I - D^{-1}A$ sum to zero, that is

$$e'L = 0 \tag{3.7}$$

Starting with $\lambda f_\alpha = Lf_\alpha$ and premultiplying both sides by *e'*, yields

$$\lambda[\sum_{x\in X} f_\alpha(x)] = \lambda(e' f_\alpha) = (e' Lf_\alpha)$$

$$= 0 \qquad \text{by (3.7)}$$

Therefore, $\lambda \neq 0$ implies $\sum_{x\in X} f_\alpha(x) = 0$, $\alpha \equiv \mu$, as required. □

We conclude the chapter by presenting an elementary landscape where $\alpha \neq \mu$.

Example 3.1: Consider a 4-city TSP where the single agent resides in city 1. Therefore, the solution space contains 6 tours, i.e.,

$$X = \begin{pmatrix} x_1 = (1,2,3,4) \\ x_2 = (1,2,4,3) \\ x_3 = (1,3,4,2) \\ x_4 = (1,3,2,4) \\ x_5 = (1,4,3,2) \\ x_6 = (1,4,2,3) \end{pmatrix}$$

Define the associated distance matrix to be $D = \begin{pmatrix} 0 & 2 & 2 & 2 \\ 2 & 0 & 2 & 1 \\ 6 & 1 & 0 & 2 \\ 6 & 2 & 3 & 0 \end{pmatrix}$

Hence the objective function vector is $f = \begin{pmatrix} f_1 = (1,2,3,4) = 12 \\ f_2 = (1,2,4,3) = 12 \\ f_3 = (1,3,4,2) = 8 \\ f_4 = (1,3,2,4) = 10 \\ f_5 = (1,4,3,2) = 8 \\ f_6 = (1,4,2,3) = 12 \end{pmatrix}$.

The arithmetic average of the f_i is μ=62/6 ≈ 10.333.

Let the normalization constant $\alpha \equiv 10 < \mu = 62/6$ yielding $f_\alpha = f - \alpha = \begin{pmatrix} +2 \\ +2 \\ -2 \\ +0 \\ -2 \\ +2 \end{pmatrix}$

Define a neighborhood N with adjacency matrix $A = \begin{pmatrix} 0 & 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 \end{pmatrix}$.

The associated Laplacian is $L = \begin{pmatrix} 1 & -1/3 & -1/3 & -1/3 & 0 & 0 \\ -1/2 & 1 & 0 & 0 & -1/2 & 0 \\ 0 & 0 & 1 & -1/3 & -1/3 & -1/3 \\ -1/2 & 0 & -1/2 & 1 & 0 & 0 \\ 0 & 0 & -1/3 & -1/3 & 1 & -1/3 \\ 0 & -1/2 & 0 & 0 & -1/2 & 1 \end{pmatrix}$.

It is easy to verify that $\mathcal{L} = (X, f_\alpha, N)$ is an elementary landscape. Since $\alpha < \mu$ this landscape provides a better upper bound on local minima then any landscape with $\alpha = \mu$.

In Chapter 4 we show that the autocorrelation spectrum of any elementary landscape (with arbitrary *T*) is consistent with an AR(1) process. We also develop four new properties of the objective function vectors associated with elementary landscapes.

Chapter 4

Arbitrary Elementary Landscape & AR(1) Process

In this chapter, we show that:

(1) any elementary landscape $\mathcal{L}$ (with arbitrary T) will possess an exponentially decaying autocorrelation spectrum consistent with an AR(1) process, i.e., $\rho_\pi(s) = \rho_\pi^{s}(1)$.

(2) any landscape $\mathcal{L}$ is elementary if a univariate time series of f_α generated from a random walk on T is consistent with an AR(1) process.

These contributions are accompanied by the proofs of four new properties of the objective function vectors associated with the elementary landscapes.

In Chapter 2.8 we presented the exponential property of the autocorrelation function of a time series created from a random walk on an elementary landscape with regular and symmetric neighborhood. Here we extend this property to arbitrary neighborhood digraphs.

4.1 Any Elementary Landscape is Consistent with an AR(1) Process

Let us propose a random walk, x_i, on a given landscape $\mathcal{L}$. This random walk yields a corresponding objective function, f_i. The function $f_{\alpha i}$ is α normalized. Weinberger (1990), defines the *sample autocovariance function* of the time series f_{α_i} of length *n*, $f_{\alpha,1}, f_{\alpha,2}, ..., f_{\alpha,n}$ generated by a random walk on *T* as

$$r_{\pi}(s) = \sum_i \pi_i \sum_j T_{ij}^{\ s} f_{\alpha i} f_{\alpha j} \tag{4.1}$$

Example 4.1: Let $\mathcal{L} = (X, f, \mathcal{N})$ be a landscape where:

$X = (x_1, x_2, x_3, x_4)$ is the solution space,

$f(x) = [f(x_1), f(x_2), f(x_3), f(x_4),] = (3,1,5,11)$ is the objective function values,

$$N = \begin{pmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 \end{pmatrix}$$ is the adjacency matrix.

Then the transition matrix is $T = \begin{pmatrix} 0 & 1/2 & 0 & 1/2 \\ 1/2 & 0 & 0 & 1/2 \\ 0 & 1/2 & 1/2 & 0 \\ 1/2 & 0 & 1/2 & 0 \end{pmatrix}$ and the corresponding steady state distribution vector is $\pi = (\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4})$. Now we propose a random walk on this landscape as:

1) Start at solution x_2 with probability $\pi_2 = \frac{1}{4}$
2) Move to solution x_1 with probability $t_{21} = \frac{1}{2}$
3) Move to solution x_4 with probability $t_{14} = \frac{1}{2}$
4) Move to solution x_3 with probability $t_{43} = \frac{1}{2}$
5) Move to solution x_3 with probability $t_{33} = \frac{1}{2}$
6) Move to solution x_2 with probability $t_{32} = \frac{1}{2}$
7) Move to solution x_1 with probability $t_{21} = \frac{1}{2}$
8) Move to solution x_2 with probability $t_{12} = \frac{1}{2}$

As a result we have the series $(f_1, f_2, f_3, f_4, f_5, f_6, f_7, f_8) =$

$= [f(x_2), f(x_1), f(x_4), f(x_3), f(x_3), f(x_2), f(x_1), f(x_2)] =$

$= (1, 3, 11, 5, 5, 1, 3, 1)$.

The sample mean of the objective function values along the simulated random walk is $\bar{x} = \frac{1+3+11+5+5+1+3+1}{10} = 3$.

Let us first calculate the lag 1 autocovariance and the variance of the f_i , $i=1,...,8$, along the random walk.

The sample autocovariance at lag 1 is

$$
\begin{aligned}
&= \pi_2 t_{21} f_\mu(x_2) f_\mu(x_1) + \pi_1 t_{14} f_\mu(x_1) f_\mu(x_4) + \pi_4 t_{43} f_\mu(x_4) f_\mu(x_3) + \\
&+ \pi_3 t_{33} f_\mu(x_3) f_\mu(x_3) + \pi_3 t_{32} f_\mu(x_3) f_\mu(x_2) + \pi_2 t_{21} f_\mu(x_2) f_\mu(x_1) + \\
&+ \pi_1 t_{12} f_\mu(x_1) f_\mu(x_2) = \\
&= \frac{1}{4}\frac{1}{2}(1-3).(3-3) + \frac{1}{4}\frac{1}{2}(3-3).(11-3) + \frac{1}{4}\frac{1}{2}(11-3).(5-3) + \\
&+ \frac{1}{4}\frac{1}{2}(5-3).(5-3) + \frac{1}{4}\frac{1}{2}(5-3).(1-3) + \frac{1}{4}\frac{1}{2}(1-3).(3-3) + \\
&+ \frac{1}{4}\frac{1}{2}(3-3).(1-3) = 2
\end{aligned}
$$

The sample variance is

$$
\begin{aligned}
&\frac{1}{(8-1)}[f_\mu(x_2) f_\mu(x_2) + f_\mu(x_1) f_\mu(x_1) + f_\mu(x_4) f_\mu(x_4) + \\
&+ f_\mu(x_3) f_\mu(x_3) + f_\mu(x_3) f_\mu(x_3) + f_\mu(x_2) f_\mu(x_2) + \\
&+ f_\mu(x_1) f_\mu(x_1) + f_\mu(x_2) f_\mu(x_2)] = \\
&= \frac{1}{7}[(1-3)^2 + (3-3)^2 + (11-3)^2 + (5-3)^2 + \\
&+ (5-3)^2 + (1-3)^2 + (3-3)^2 + (1-3)^2] = \frac{84}{7} = 12
\end{aligned}
$$

Thus, the sample autocorrelation function at lag 1 is the ratio of the sample autocovariance at lag 1 divided by the sample variance which equals one-sixth.

In Section 2.7 we discussed that, by assuming a regular and symmetric neighborhood, Weinberger (1990) and Weinberger and Stadler (1993) consider a doubly stochastic transition matrix T with uniform steady state vector π. Ross (1997) shows that the steady state distribution of such a transition matrix is

$$\pi = \left[\frac{1}{|X|} \right]$$

This assumption and Theorems 3.2 and 3.7 from Chapter 3 lead to the conclusion that the normalization constant for symmetric and regular T is

$$\alpha \equiv \mu$$

The theoretical autocorrelation function for *any* $\mathcal{L}$ can be written as (Stadler (1996))

$$\rho_{\pi}(s) = \frac{f_{\alpha}^{'} \Pi T^{s} f_{\alpha}}{f_{\alpha}^{'} \Pi f_{\alpha}}, \tag{4.2}$$

where Π is a diagonal matrix with elements $\Pi_{ii} = \pi_i$. The doubly stochastic matrix T includes the symmetric-regular T (symmetric T means $t_{ij}=t_{ji}$ for every i and j, regular T means the sum of all rows and column sums to one) as a special case. Since $\pi = [\frac{1}{|X|}]$ is true if and only if T is doubly stochastic, as proven by Ross (1997), the theoretical autocorrelation function for a landscape $\mathcal{L}$ with doubly stochastic T is

$$\rho_{\pi}(s) = \frac{f_{\alpha}^{'} T^{s} f_{\alpha}}{f_{\alpha}^{'} f_{\alpha}} \tag{4.3}$$

If T is symmetric-regular, then $\mathcal{L}$ is elementary if and *only* if a univariate time series generated from a random walk on T is consistent with an autoregressive process of order 1, i.e., an AR(1) process (Stadler (1996)). For a symmetric-regular T, the set of f_α associated with an elementary $\mathcal{L}$ is identical to the set of f_α from a random walk on T that are consistent with an AR(1) process (Stadler (1996), Weinberger (1990), Stadler, Seitz and Wagner (2000), Box and Jenkins (1976)).

Instead of using Equation 4.3 as a characterization of the landscape one may also use the sample autocorrelation function as defined in Equation 4.1, (Stadler, Happel (1992)). Let us now consider an *arbitrary* transition matrix *T,* i.e. , we will not assume that T is symmetric or regular.

Theorem 4.1: For any elementary landscape $\mathcal{L}$, the autocorrelation function is $\rho_\pi(s) = (1-\lambda)^S$ where λ is an eigenvalue of the Laplacian, L, and $\rho_\pi(s) = \rho_\pi^{\,s}(1)$.

Proof: As was shown in Section 3.1, a landscape $\mathcal{L}$ is elementary if and only if $Tf_\alpha = (1-\lambda)\, f_\alpha$. First note that $T^s f_\alpha = (1-\lambda)^s f_\alpha$. Substituting this relation into Equation 4.2, and simplifying yields the autocorrelation function

$$\rho_\pi(s) = \frac{f_\alpha^{'} \Pi T^s f_\alpha}{f_\alpha^{'} \Pi f_\alpha} = \frac{f_\alpha^{'} \Pi (1-\lambda)^s f_\alpha}{f_\alpha^{'} \Pi f_\alpha} = (1-\lambda)^s \tag{4.4}$$

For $s=1$, $\rho_\pi(1) = (1-\lambda)$, which imples that the autocorrelation function can be written as the exponential function $\rho_\pi(s) = \rho_\pi^{\,s}(1)$. This is a fundamental characteristic of AR(1) processes, (Box and Jenkins (1976)). □

Theorem 4.2: Any landscape $\mathcal{L}$ which generates an AR(1) time series by means of a random walk on T is elementary.

Proof: For all x_i, $i=1,\ldots, n,$ visited by the random walk on T with length n, denote by $\alpha = \pi Tf = \pi f = E(f_i)$. Now fix the index i. Let $[z_t] = [f_i - \alpha]$ be the normalized time series yielded from the random walk, i.e., z_t is the t^{th} normalized solution visited in the generation of the time series with value $f_i - \alpha = f_{\alpha_i}$. In classical time series, see Box and Jenkins (1976), a theoretical AR(1) process is defined by the recurrence equation,

$$z_t = \phi_1 z_{t-1} + a_t \, ,$$

which consists of a sequence of uncorrelated random variables with mean zero and constant variance, and a coefficient $\phi_1 = \rho(1)$.

Consider the temporally adjacent pair, z_t and $z_{t-1.}$ By definition, z_{t-1} can be written as $z_{t-1} \equiv f_{\alpha i}$ for *some* fixed i. For a given z_{t-1} (corresponding to an x_i), there are d_i possible values of z_t associated with the neighbors of z_{t-1}. Let us write them as $z_{t,j}$, $j = 1,\ldots,d_i$. In our random walk over a given finite $\mathcal{L}$, the parameters a_t are random deviates drawn from a finite discrete probability distribution with an expected value of zero and constant variance. This follows directly from the fact that $a_t = z_t - \phi_1 z_{t-1}$ and $E(z_t) = 0$ for all t.

For any neighbor $z_{t,j}$ of z_t, it is true that

$$z_{t,j} = \phi_1 z_{t-1} + a_{t,j} \tag{4.5}$$

Summing Equation 4.5 over all possible values of $z_{t,j}$ where $j = 1,\ldots,d_i$ (all the neighboring solutions of $z_{t\text{-}1}$) and averaging yields

$$\sum_{j=1}^{d_i} \frac{z_{t,j}}{d_i} = \phi_1 \sum_{j=1}^{d_i} \frac{z_{t-1}}{d_i} + \sum_{j=1}^{d_i} \frac{a_{t,j}}{d_i} \tag{4.6}$$

Notice that

$$\sum_{j=1}^{d_i} \frac{z_{t,j}}{d_i} = \underset{x_j \in \mathcal{N}_i}{Avg}\, f_{\alpha j} = T_i f_\alpha \tag{4.7}$$

and

$$\sum_{j=1}^{d_i} \frac{z_{t-1}}{d_i} = f_{\alpha i} \tag{4.8}$$

Substituting Equation 4.7 and Equation 4.8 into Equation 4.6 we obtain

$$T_i f_\alpha = \phi_1 f_{\alpha i} + \sum_{j=1}^{d_i} \frac{a_{t,j}}{d_i},\ i=1,\ \ldots,/X/ \tag{4.9}$$

Taking the expectation of Equation 4.9 yields (because E($a_{t,j}$)=0)

$$T_i f_\alpha = \phi_1 f_{\alpha i}, \quad i=1, \ldots, |X|. \tag{4.10}$$

In matrix form, Equation 4.10 can be written as $Tf_\alpha = \phi_1 f_\alpha$, which implies that the landscape $\mathcal{L}$ associated with T is elementary with eigenvalue
$\lambda = 1-\phi_1 = 1-\rho(1)$. □

Theorem 4.1 and Theorem 4.2 show that for any *arbitrary* T, a landscape $\mathcal{L}$ is elementary if and only if a univariate time series on f_α generated from a random walk on T is consistent with an autoregressive process of order 1, i.e. the associated autocorrelation function will be governed by an exponential function, the AR(1) autocorrelation function.

4.2 Orthogonality Properties of Elementary Landscapes

This section presents some properties of the objective function vector of elementary landscapes with respect to a steady state probability vector of the transition matrix.

Theorem 4.3: For *any* non flat landscape $\mathcal{L}$ (with arbitrary transition matrix T), all f_α yielding elementary $\mathcal{L}$ are orthogonal to the steady state vector, π, of T, i.e., $f_\alpha \perp \pi$.

Proof: Since $\mathcal{L}$ is elementary, Equation 3.2 holds which implies that $Tf_\alpha = (1-\lambda) f_\alpha$. Left multiplication by π' yields $\pi' T f_\alpha = (1-\lambda)\pi' f_\alpha$. Since $\pi' T = \pi$ (Kemeny and Snell (1960)), then $\pi' f_\alpha = (1-\lambda)\pi' f_\alpha$ or equivalently, $\lambda\pi' f_\alpha = 0$. Excluding flat landscapes ,where $\lambda=0$, we obtain $\pi' f_\alpha = 0$, i.e., $f_\alpha \perp \pi$. Hence, the weighted sum of the normed f_{α_i} is zero and any objective function vector associated with an elementary landscape is perpendicular to π.

For a case of flat landscape when $\lambda=0$, and normalizing it by the expectation α, we get $\pi' f_\alpha = \pi'(f-\alpha) = \pi' f - \pi'\alpha = \alpha - \alpha = 0$. Hence $f_\alpha \perp \pi$. □

Theorem 4.4: **F**or any landscape $\mathcal{L}$ with arbitrary *aperiodic* transition matrix T, all f_α, for which a random walk on T is consistent with an AR(1) process, are orthogonal to the steady state vector π.

Proof: Let L be a landscape for which a random walk is consistent with an *AR(1)* process. For any AR(1) process, $\rho_\pi(s) = \rho_\pi^{s}(1)$. Substituting into the matrix form

$$\rho_\pi(s) = \frac{f_\alpha' \Pi T^s f_\alpha}{f_\alpha' \Pi f_\alpha} \tag{4.11}$$

of the theoretical autocorrelation function we obtain

$$f_\alpha' \Pi T^s f_\alpha = \rho_\pi^{s}(1) f_\alpha' \Pi f_\alpha, \tag{4.12}$$

i.e., $f_\alpha' \Pi T^s f_\alpha$ is linearly proportional to $\rho_\pi^s(1)$. If $|\rho_\pi(1)| < 1$, $\lim_{s\to\infty} \rho_\pi^s(1) = 0$.

Hence

$$\lim_{s\to\infty} f_\alpha' \Pi T^s f_\alpha = 0. \tag{4.13}$$

Define G to be a square matrix where $G_{ij} = 1 \; \forall \, i,j$, then $\lim_{s\to\infty} T^s = G\Pi$, see Kulkarni (1995).

Since matrix multiplication is associative, Equation 4.13 can be expressed as

$$0 = \lim_{s\to\infty} f_\alpha' \Pi T^s f_\alpha = f_\alpha' \Pi G \Pi f_\alpha = (\Pi f_\alpha)' G (\Pi f_\alpha)$$

Let Πf_α denote the following vector

$$\Pi f_\alpha = (\pi_1 f_{\alpha,1}, \pi_2 f_{\alpha,2}, ..., \pi_{|X|} f_{\alpha,|X|})'$$

and let

$$G(\Pi f_\alpha) = (s, s, ..., s)' = s.e,$$

where s is defined as

$$\text{s} = \pi_1 f_{\alpha,1} + \pi_2 f_{\alpha,2} + ... + \pi_{|X|} f_{\alpha,|X|} = \pi' f_\alpha$$

and $e = (1,1,...,1)'$. It follows that

$$\lim_{s\to\infty} f_\alpha' \Pi T^s f_\alpha = f_\alpha' \Pi G \Pi f_\alpha = (\Pi f_\alpha)' G (\Pi f_\alpha) =$$

$$= (\Pi f_\alpha)' s.e \; = \; s((\Pi f_\alpha)'.e) = s^2 = (\pi' f_\alpha)^2 .$$

Therefore $0 = \lim_{s \to \infty} f_\alpha' \Pi T^s f_\alpha = (\pi' f_\alpha)^2$

which is true if and only if $\pi' f_\alpha = 0$. Therefore, $f_\alpha \perp \pi$. $\square$

Theorem 4.5: Let T be an arbitrary transition matrix for which $\pi' T f_\alpha = \pi' f_\alpha = 0$, then $\alpha = \pi' f$.

Proof: $\pi' T f_\alpha = 0$ implies $\pi' T f - \pi' T \alpha = 0$ which yields $\pi' f - \pi' \alpha = 0$.
Thus $\pi' f = \pi' \alpha = \alpha$ and $\alpha = \pi' f$. $\square$

Theorem 4.6: For doubly stochastic T, then $\pi' f = \mu$.

Proof: For doubly stochastic T, $\pi = [\frac{1}{|X|}]$. Hence $\pi' f = \frac{\sum_{\forall i \in X} f_i}{|X|} = \mu$. $\square$

In Chapter 3 we proved that if a normalization constant exists such that the normalized objective function is an eigenvector of the Laplacian, then this constant is unique. Here we prove that if the landscape is elementary, then the normalization constant is the expectation, α, of the objective function given T. These two statements imply that every elementary landscape has an objective function normalized by the expectation α.

In Chapter 5 we continue with the development of a method to obtain elementary landscapes using the composition of elementary landscapes, analyze the character of the new elementary landscape (smooth or rugged) with respect to the building elementary landscapes, and calculate lower (upper) bound of a local minimum (maximum) for the elementary landscapes.

Chapter 5

Some Additional Properties of Elementary Landscapes

The contributions of this chapter are:

(1) composition of elementary landscapes is elementary landscape.

(2) develop the first non trivial lower (upper) bound of a local minimum (maximum) for the elementary landscapes.

In addition we analyze the character of the composite elementary landscape (smooth or rugged) with respect of the building elementary landscapes.

5.1. Composite Elementary Landscapes

Let $\mathcal{L}_A = (X, f, N_A)$ and $\mathcal{L}_B = (X, f, N_B)$ be two landscapes differing *only* by their neighborhood definitions. Suppose that $\mathcal{L}_B$ is the *composite* landscape generated by performing two sequential moves defined by N_A , i.e., N_B has associated transition matrix $T_B = T_A T_A = T_A^2$. In general, $\mathcal{L}_B$ could be result of an arbitrary number, *n*, of sequential moves under N_A, i.e., N_B would have associated transition matrix $T_B = T_A^n$. We termed such a neighborhood as an "*n*-step neighborhood." However, from the perspective of $\mathcal{L}_B$, such an N_B is simply an alternative "one-step" neighborhood. Similarly, a new composite landscape, $\mathcal{L}_C$, could be generated by first performing a move according to N_A and then according to N_B, yielding $T_C = T_A T_B$ (Kemeny and Snell (1960)).

Let $\mathcal{L}_i = (X, f, N_i)$ be the i-step landscape relative to L with a transition matrix T^i. Let $[\, z_t^i \,] = [f_{j,t} - \alpha] = [f_{\alpha(j,t)}\,]$ be the time series yielded from a random walk, such as discussed in Section 2.7, on $\mathcal{L}_i$, i.e., starting at any $x_j \in X$, the t^{th} normalized solution visited in the generation of the time series based on the neighborhood N_i has value, z_t^i.

Classically an AR(1) process is defined by the recurrence equation, Equation 2.6, where $p=1$

$$z_t = \phi z_{t-1} + a_t \tag{5.1}$$

Suppose that $\mathcal{L}_i = (X, f, N_i)$ is elementary. The time series $[\, z_t^i \,]$ associated with $\mathcal{L}_i = (X, f, N_i)$ can be written as

$$z_t^i = \psi z_{t-1}^i + a_t^i \tag{5.2}$$

Theorem 5.1: For any elementary $\mathcal{L}$ (given a real eigenvalue of the Laplacian) with connected N, the corresponding 2-step neighborhood yields a smooth elementary landscape $\mathcal{L}_2$.

Proof: By Theorem 4.1 if $\mathcal{L}$ is an elementary landscape, its associated autocorrelation function is consistent with an AR(1) process, i.e., $\rho(s) = \rho^s(1)$. Let us consider the two one-step recurrence equations for two steps neighborhood

$$z_{t-1} = \phi z_{t-2} + a_{t-1} \tag{5.3}$$

$$z_t = \phi z_{t-1} + a_t \tag{5.4}$$

Substituting (5.3) in (5.4) yields:

$$z_t = \phi(\phi z_{t-2} + a_{t-1}) + a_t = \phi^2 z_{t-2} + \phi a_{t-1} + a_t \tag{5.5}$$

In comparing Equations 5.1 and 5.5 where i = 2, we observe that

$$z_\tau^2 = z_t\,, \qquad z_{\tau-1}^2 = z_{t-2}\,, \qquad \psi = \phi^2 \qquad \text{and} \qquad a_\tau^2 = \phi\, a_{t-1} + a_{t.}$$

The next term in z_τ^2 is

$$z_{\tau+1}^2 = z_{t+2} = \phi(\phi z_t + a_{t+1}) + a_{t+2} = \phi^2 z_t + \phi a_{t+1} + a_{t+2} = \psi z_\tau^2 + a_{\tau+1}^2 \tag{5.6}$$

The result establishes that the a_τ^2 are uncorrelated with expected value zero and variance equal to $(\phi^2 + 1)\sigma^2$. This fact and the form of the recurrence equation for z_τ^2 by Theorem 4.2 are sufficient to establish that the composite landscape, $\mathcal{L}_2$, resulting from two sequential moves under N is elementary. Since $\mathcal{L}_2$ is elementary,

$$T^2 f_\alpha = \psi f_\alpha,$$

that implies

$$(I - T^2) f_\alpha = (1 - \psi) f_\alpha$$

Hence

$$L_2 f\alpha = \lambda_2 f\alpha \qquad (5.7)$$

where T^2 is the transition matrix of $\mathcal{L}_2$. L_2 is the Laplacian matrix associated with N_2, and L_2 has eigenvalue λ_2. Furthermore

$$\lambda_2 = 1 - \psi = 1 - \phi^2 = 1 - (1 - \lambda_1)^2 = 2\lambda_1 - \lambda_1^2$$

where λ_1 is the eigenvalue associated with $\mathcal{L}$. Since $0 \le \lambda_1 \le 2$ then $0 \le \lambda_2 \le 1$. Hence a two-step neighborhood associated with any elementary $\mathcal{L}$ will always yield a smooth elementary landscape $\mathcal{L}_2$. □

Theorem 5.1 considered the composition of two identical landscapes. We will now consider the composition of two elementary landscapes differing only in their neighborhood definition and prove similar statements.

Theorem 5.2: Let $\mathcal{L}_A = (X, f, N_A)$ and $\mathcal{L}_B = (X, f, N_B)$ be two elementary landscapes. The composite neighborhood, $\mathcal{L}_C = (X, f, N_C)$, generated by a move from N_A followed by a move from N_B is an elementary landscape.

Proof: Both $\mathcal{L}_A$ and $\mathcal{L}_B$ are consistent with an AR(1) process. The corresponding recurrence equations are

$$z_t = \phi_A z_{t-1} + a_t \qquad \text{and} \qquad z_{t+1} = \phi_B z_t + a_{t+1}.$$

The recurrence equation for the composite random walk is

$$z_{t+1} = \phi_B(\phi_A z_{t-1} + a_t) + a_{t+1} = \phi_B\phi_A z_{t-1} + \phi_A a_t + a_{t+1} = \omega z_{t-1} + c_{t+1} \qquad (5.8)$$

which, analogous to Equation 5.6 yields $z_{\tau+1} = \omega\, z_\tau + c_{\tau+1}$.

Therefore, since the c_τ are uncorrelated with expected value zero and variance equal to $(\phi_A^2 + 1)\sigma^2$, we conclude that the composite landscape, $\mathcal{L}_C$, is elementary. $\square$

Theorem 5.2 may be used to determine the required properties of $\mathcal{L}_A$ and $\mathcal{L}_B$ to cause $\mathcal{L}_C$ to be either smooth or rugged:

(1) if $\mathcal{L}_A$ and $\mathcal{L}_B$ are smooth elementary landscapes ($0 \le \phi_{A,} \phi_B \le 1$) then $\mathcal{L}_C$ is smooth ($0 \le \omega \le 1$).

(2) if $\mathcal{L}_A$ and $\mathcal{L}_B$ are rugged elementary landscapes ($-1 \le \phi_A \le 0$ and $-1 \le \phi_B \le 0$), then $\mathcal{L}_C$ is smooth ($0 \le \omega \le 1$).

(3) Otherwise, $\mathcal{L}_C$ is rugged. For example, if $\mathcal{L}_A$ is smooth and $\mathcal{L}_B$ is rugged ($0 \le \phi_A \le 1$ and $-1 \le \phi_B \le 0$), then $\mathcal{L}_C$ is rugged ($-1 \le \omega \le 0$). (The same result follows when $\mathcal{L}_B$ is smooth and $\mathcal{L}_A$ is rugged.)

Corollary 5.1: Let $\mathcal{L}_i$, $i=1,\ldots,n$ be a set of elementary landscapes differing only in their neighborhood definitions and define the Λ_φ, j $=1,\ldots,n$ to be a *chain* of composite landscapes such that $\Lambda_\varphi = \overset{j}{\underset{i=1}{\oplus}}\{\mathcal{L}_i\}$ where $\overset{j}{\underset{i=1}{\oplus}}$ implies the composition of the $\mathcal{L}_i$, in order, for $i = 1, \ldots, j$. Each of the Λ_φ is elementary landscape.

Proof: Follows directly from Theorem 5.2. □

5.2. Lower Bounds on Local Minima and Upper Bounds on Local Maxima for Elementary Landscapes

In Theorem 3.3 we prove that all local minima (maxima) of an elementary landscape are bounded above (below) by α. Using the information stability parameter discussed in Chapter 2.9, we now derive a *lower* (*upper*) bound on local minima (maxima) for any elementary landscape.

Theorem 5.3: Local minima of elementary landscapes are bounded below by $\alpha - \frac{\varepsilon^*}{\lambda}$, where λ is the eigenvalue of the associated L.

Proof: Let x_i^* be a local minimum, $f_{\alpha\, i,\, max}$, be the maximum value of f_α in N(x_i) and $Avg_{y\in\mathcal{N}(x_i)} f_\alpha(y)$ be the average value of f_α in $N(x_i)$. For any local minimum, $\left| f_\alpha(x_i^*) - f_{\alpha i,\max} \right| \leq \varepsilon^*$ for a time series of sufficient length.

Hence,

$$Avg_{y\in\mathcal{N}(x_i)} f_\alpha(y) \leq f_{\alpha i,\max} \tag{5.9}$$

and

$$f_\alpha(x_i^*) - f_{\alpha i,max} \leq f_\alpha(x_i^*) - Avg_{y\in\mathcal{N}(x_i)} f_\alpha(y)$$

which implies

$$-\varepsilon^* \leq f_\alpha(x_i^*) - f_{\alpha i,max} \leq f_\alpha(x_i^*) - Avg_{y\in\mathcal{N}(x_i)} f_\alpha(y) \tag{5.10}$$

and therefore,

$$-\varepsilon^* \leq f_\alpha(x_i^*) - T_i f_\alpha \tag{5.11}$$

Since the landscape is elementary by Equation 3.2 $T_i f_\alpha = (1-\lambda) f_\alpha(x_i)$. Substitution into Equation 5.11 yields $-\varepsilon^* \leq f_\alpha(x_i^*) - (1-\lambda) f_\alpha(x_i^*)$, where $0 \leq \lambda \leq 2$. Therefore $-\dfrac{\varepsilon^*}{\lambda} \leq f_\alpha(x_i^*)$ which implies $\alpha - \dfrac{\varepsilon^*}{\lambda} \leq f(x_i^*)$. □

Theorem 5.4: Local maxima for elementary landscapes are bounded above by $\alpha + \frac{\varepsilon^*}{\lambda}$.

Proof: Analogical of the proof of Theorem 5.3. $\square$

In the next Chapter 6 we extend the Laplacian equation to *General Laplacian Equation*. We prove that landscapes satisfied such equation if and only if an univariate time series generated from a random walk is consistent with AR(p) process and extend the concept for favorable neighborhood for a local search.

Chapter 6

The Characteristic Landscape Equation

For an AR(p) Landscapes

The primary contributions of this chapter are:

(1) developing the characteristic landscape equation for an AR(2) and AR(p) landscapes.

(2) proving that a landscape $\mathcal{L}$ is satisfied the corresponding characteristic landscape equation if and *only* if a univariate time series on f_α generated from a random walk on T is consistent with an AR(p) process.

(3) providing some landscapes (not elementary) favorable for local search.

Investigating the properties of AR(2) landscapes we separate a set of landscapes which under some additional conditions about the parameters φ_1 and φ_2 of stationary AR(2) processes, have properties favorable for local search. We also extend the concept of "favorable neighborhood" for a local search.

Let α be the expected value of f (Theorem 4.5) and let $[z_t]= [f_{i,t}\ -\ \alpha] = [\ f_{\alpha\ i}\ ]$ be the time series yielded from a random walk on $\mathcal{L}$ starting at x_i, i.e., the t^{th} normalized solution visited in the generation of the time series has objective function, $z_t = f_{\alpha\ i}$. As we discussed in Section 2.8, an AR(2) process is defined by the recurrence equation,

$$z_t = \phi_1\ z_{t-1} + \phi_2\ z_{t-2} + a_t \tag{6.1}$$

Consider the temporally adjacent triplet, z_t, z_{t-1} and z_{t-2}. By definition, $z_{t-2} \equiv f_{\alpha\ i}$ for *some* x_i. For the specified z_{t-2}, the d_i neighbors of x_i are the $x_j \in \mathcal{N}_i$ with values $z_{t-1,j}$. For a given $x_j \in \mathcal{N}_i$, the d_j neighbors are the $x_k \in \mathcal{N}_j$ with values $z_{t,k} \equiv f_{\alpha,\mathrm{k}}$.

In a random walk over a given finite $\mathcal{L}$, the parameters a_t are random deviates drawn from a finite discrete probability distribution with an expected value of zero. This follows directly from the fact that $E(z_t) = 0$ for all t and $a_t = z_t - \phi_1\ z_{t-1} - \phi_2\ z_{t-2}$.

6.1. The Characteristic Landscape Equation for AR(2) Landscapes

In Chapter 2.6 we have a discussion about the elementary landscapes, define based on the Grover's wave equation (2.1) and their properties, which make such landscapes favorable for a local search. Here we extend the knowledge about the landscapes developing a characteristics neighborhood equation for AR(2) process and prove the equivalency: the landscape satisfies the characteristic equation neighborhood equation for AR(2) and the autocorrelation function of the corresponding time series based on a random walk is associated with AR(2) process. Such landscapes we call AR(2) landscapes. Later on based on this equivalency we investigate some properties of AR(2) landscapes and define some landscapes that are not elementary but have the same favorable for a local search properties.

Theorem 6.1: If the time series based on a random walk on the landscape is consistent with an AR(2) process, then the landscape satisfies the characteristic landscape equation, $(T^2 - \phi_1 T) f_\alpha = \phi_2 f_\alpha$.

Proof: The proof proceeds in three steps:

(1) the one step neighborhood of a specific neighbor of solution x_j which is a specific neighbor of x_i is first considered;

(2) the results are expanded to consider all neighbors, $x_j \in \mathcal{N}_i$;

(3) all possible starting solutions, x_i, are considered.

Step 1: We are given a specific solution x_i, with associated value, $z_{t-2} \equiv f_{\alpha\, i}$. Consider a specific one step neighbor of x_i, $x_j \in \mathcal{N}_i$, with associated value $z_{t-1,j} \equiv f_{\alpha,\, j}$. The cardinality of $\mathcal{N}_j$ is d_j. With this fixed x_i and x_j, we average Equation 6.1 over the d_j neighbors of x_j which yields (reading $k \in \mathcal{N}_j$ as "k such that $x_k \in \mathcal{N}_j$")

$$\sum_{k\in\mathcal{N}_j}\frac{z_{t,k}}{d_j}=\phi_1\sum_{k\in\mathcal{N}_j}\frac{z_{t-1,j}}{d_j}+\phi_2\sum_{k\in\mathcal{N}_j}\frac{z_{t-2}}{d_j}+\sum_{k\in\mathcal{N}_j}\frac{a_{t,k}}{d_j} \tag{6.2}$$

Observe that

$$\sum_{k\in\mathcal{N}_j}\frac{z_{t,k}}{d_j}=\underset{k\in\mathcal{N}_j}{Avg}\, f_{\alpha,k}=T_j f_\alpha,\quad \sum_{k\in\mathcal{N}_j}\frac{z_{t-1,j}}{d_j}=f_{\alpha,j} \text{ and } \sum_{k\in\mathcal{N}_j}\frac{z_{t-2}}{d_j}=f_{\alpha,i}$$

where T_j is the j^{th} row of *T*. Substituting into Equation 6.2 we obtain

$$T_j f_\alpha=\phi_1 f_{\alpha,j}+\phi_2 f_{\alpha,i}+\sum_{k\in\mathcal{N}_j}\frac{a_{t,k}}{d_j},\ \ i=1,\ \ldots,/X/,\ j\colon x_j\in\mathcal{N}_i \tag{6.3}$$

Step 2: Taking the average of Equation 6.3 over the d_i neighbors of x_i, i.e., over the $x_j \in \mathcal{N}_i$, we obtain

$$\frac{1}{d_i}\sum_{j\in\mathcal{N}_i}T_j f_\alpha=\phi_1\frac{1}{d_i}\sum_{j\in\mathcal{N}_i}f_{\alpha,j}+\phi_2\frac{1}{d_i}\sum_{j\in\mathcal{N}_i}f_{\alpha,i}+\frac{1}{d_i}\sum_{j\in\mathcal{N}_i}\sum_{k\in\mathcal{N}_j}\frac{a_{t,k,j}}{d_j},\ \ i=1,\ \ldots,/X/ \tag{6.4}$$

Observing that $\frac{1}{d_i}\sum_{j\in\mathcal{N}_i}T_j f_\alpha=T_i^2 f_\alpha$, $\frac{1}{d_i}\sum_{j\in\mathcal{N}_i}f_{\alpha,j}=T_i f_\alpha$, and $\frac{1}{d_i}\sum_{j\in\mathcal{N}_i}f_{\alpha,i}=f_{\alpha,i}$ and substituting into Equation 6.4 yields

$$T_i^2 f_\alpha=\phi_1 T_i f_\alpha+\phi_2 f_{\alpha,i}+\frac{1}{d_i}\sum_{j\in\mathcal{N}_i}\sum_{k\in\mathcal{N}_j}\frac{a_{t,k,j}}{d_j} \tag{6.5}$$

Because the error terms have an expectation zero, taking the expectation of Equation 6.5 yields

$$T_i^2 f_\alpha = \phi_1 T_i f_\alpha + \phi_2 f_{\alpha,i}\,, \quad i=1, \ldots, |X| \tag{6.6}$$

Step 3: In matrix form, Equation 6.6 can be written as

$$T^2 f_\alpha = \phi_1 T f_\alpha + \phi_2 f_\alpha$$

or

$$(T^2 - \phi_1 T) f_\alpha = \phi_2 f_\alpha \tag{6.7}$$

This completes the proof. □

Equation 6.6 is the c*haracteristic landscape equation* for an AR(2) process. If $\phi_2 =$ 0, Equation 6.7 degenerates to the classical Laplacian equation for an AR(1) process, i.e., if $\phi_2 = 0$ and if T is invertible, Equation 6.7 may be expressed as $Tf_\alpha = \phi_1 f_\alpha$. If $\mathcal{G} = T^2 - \phi_1 T$ is a stochastic matrix then an equivalent classical elementary landscape is present with associated Laplacian equation, $\mathcal{G} f_\alpha = \phi_2 f_\alpha$.

Let $[\, f_{\alpha_i} \,]$ be a time series of length n generated by a random walk on T. In Chapter 4 we talk about the Weinberger (1990) definition of the sample autocorrelation function which is defined by Equation 4.1. We show that the matrix form of the theoretical autocorrelation function for any $\mathcal{L}$ is Equation 4.2

$$\rho_\pi(s) = \frac{f_\alpha^{'} \Pi T^s f_\alpha}{f_\alpha^{'} \Pi f_\alpha}$$

Theorem 6.2: If the landscape satisfies the equation $(T^2 - \phi_1 T) f_\alpha = \phi_2 f_\alpha$, then the time series based on a random walk on this landscape is consistent with an AR(2) process.

Proof: Using Box and Jenkins (1976) approach to analyze a time series, we will prove that behavior of the theoretical autocorrelation function $\rho_\pi(s)$ of such a time series is consistent with the autocorrelation function of an AR(2) process.
We first write the Equation 6.7 in the form

$$T^2 f_\alpha = \phi_1 T f_\alpha + \phi_2 f_\alpha \tag{6.8}$$

Further, noting that $T^0 = I$

$$\rho_\pi(0) = \frac{f_\alpha^{'} \Pi T^0 f_\alpha}{f_\alpha^{'} \Pi f_\alpha} = 1 \tag{6.9}$$

Premultiplying Equation 6.8 by T^{-1} and substituting in Equation 4.2 yields

$$\rho_\pi(1) = \frac{f_\alpha^{'} \Pi T f_\alpha}{f_\alpha^{'} \Pi f_\alpha} = \frac{f_\alpha^{'} \Pi \phi_1 f_\alpha + f_\alpha^{'} \Pi \phi_2 T^{-1} f_\alpha}{f_\alpha^{'} \Pi f_\alpha} = \phi_1 \rho_\pi(0) + \phi_2 \rho_\pi(-1) \tag{6.10}$$

Similarly,

$$\rho_\pi(2) = \frac{f_\alpha' \Pi T^2 f_\alpha}{f_\alpha' \Pi f_\alpha} = \frac{f_\alpha' \Pi \left(\phi_1 T f_\alpha + \phi_2 f_\alpha\right)}{f_\alpha' \Pi f_\alpha} = \phi_1 \rho_\pi(1) + \phi_2 \rho_\pi(0) \tag{6.11}$$

The general relation is obtained by premultiplying Equation 6.8 by T^{n-2} and substituting $T^n f_\alpha$ into Equation 4.2 yields the recurrent formula for the autocorrelation function of such landscape,

$$\rho_\pi(n) = \phi_1 \rho_\pi(n-1) + \phi_2 \rho_\pi(n-2) \tag{6.12}$$

Observing that $\rho_\pi(-n) = \rho_\pi(n)$, the simultaneous solution of Equations 6.9, 6.10 and 6.11 yields

$$\rho_\pi(1) = \frac{\phi_1}{1-\phi_2} \quad \text{and} \quad \rho_\pi(2) = \phi_2 + \frac{\phi_1^2}{1-\phi_2}$$

which corresponds to the classical autocorrelation function of an AR(2) process as discussed in Box and Jenkins (1976). □

6.2. Properties of an AR(2) Landscapes

In this section we investigate the structure of an AR(2) landscape from the perspective of local minima using the already developed characteristic landscape equation 6.7 for AR(2) landscapes. This new development extends the set of elementary landscapes to more general set of landscapes favorable for a local search.

6.2.1 An Upper Bound for Local Minima for the Two-Step Neighborhood

Consider a landscape $\mathcal{L}$ consistent with an AR(2) process. For any solution x_i and hence for any two-step neighborhood local minimum, x_i^{**}, the following relation is true:

$$T_i^2 f_\alpha - \phi_1 T_i f_\alpha = \phi_2 f_{\alpha i}^{**} \tag{6.13}$$

The average of all two step neighbors of x_i^{**}, $\underset{k \in \mathcal{N}^2}{Avg} f_{\alpha,k}$, is equal to $T_i^2 f_\alpha$. Since $f_{\alpha i}^{**}$ is less than or equal to the objective function value for any two step neighbor,

$$f_{\alpha i}^{**} \leq T_i^2 f_\alpha \tag{6.14}$$

Equations 6.13 and 6.14 imply that $f_{\alpha,i}^{**} - \phi_1 T_i f_\alpha \leq \phi_2 f_{\alpha,i}^{**}$ which yields

$$(1-\phi_2) f_{\alpha,i}^{**} \leq \phi_1 T_i f_\alpha \tag{6.15}$$

Here we limit attention to stationary AR(2) processes where $0 < 1-\phi_2 < 2$ (Box and Jenkins (1976)). Since $1-\phi_2 > 0$,

$$f_{\alpha i}^{**} \le \frac{\phi_1}{1-\phi_2} T_i f_\alpha \tag{6.16}$$

Hence $f_{\alpha i}^{**}$ is bounded above by $\rho_\pi(1)T_i f_\alpha$ which is equivalent to the autocorrelation at lag *1* times the average objective function value for the one step neighbors of x_i^{**}. This states that for an AR(2) landscapes any two step local minimum is bounded by $\rho_\pi(1)T_i f_\alpha$. Hence for some landscapes the two step local minimum will be better then the one step local minimum.

6.2.2 An Upper Bound for Local Minima for the One-Step Neighborhood

For an AR(2) landscape, the set of all one-step local minima, the x_i^*, fall into two subsets:

M_1 consists of x_i^* that are also two-step local minimum

M_2 consists of x_i^* that are *not* two-step local minimum

First, we will consider M_1. Any $x_i^* \in M_1$ with value $f_{\alpha i}^*$ satisfies Equation 6.13 and the following relations also hold:

$$f_{\alpha i}^* \le T_i f_\alpha \tag{6.17}$$

$$f_{\alpha i}^* \le T_i^2 f_\alpha \tag{6.18}$$

By Equation 6.16,

$$f_{\alpha i}^{*} \le \frac{\phi_1}{1-\phi_2} T_i f_\alpha \tag{6.19}$$

We must consider two cases: $\phi_1 \ge 0$ and $\phi_1 \le 0$.

- If $\phi_1 \ge 0$, multiplying Equation 6.17 by $-\phi_1$ and adding $T_i^2 f_\alpha$ to both sides yields $T_i^2 f_\alpha - \phi_1 f_{\alpha i}^{*} \ge T_i^2 f_\alpha - \phi_1 T_i f_\alpha = \phi_2 f_{\alpha i}^{*}$ which implies $\left(\phi_1 + \phi_2\right) f_{\alpha i}^{*} \le T_i^2 f_\alpha$.

This leads to the following two subcases where we assume that $T_i^2 f_\alpha \le 0$. (If $T_i^2 f_\alpha > 0$, no additional meaningful conclusions can be drawn.):

(i) If $\phi_1 + \phi_2 > 0$ and $T_i^2 f_\alpha \le 0$, $f_{\alpha,i}^{*} \le 0$ and $f_i^{*} \le \alpha$, i.e., arbitrarily poor local optima of class M_1 cannot exist

(ii) If $\phi_1 + \phi_2 < 0$ and $T_i^2 f_\alpha \le 0$, $f_{\alpha,i}^{*} \ge 0$. However, in this case, Equation (16) requires that $f_{\alpha,i}^{*} \le 0$. Therefore $f_{\alpha,i}^{*} = 0$ or $f_i^{*} = \alpha$. This implies that all $x_i \in M_1$ have $f_i^{*} = \alpha$.

- If $\phi_1 \le 0$, multiplying Equation 6.17 by $\dfrac{\phi_1}{1-\phi_2}$ yields $\dfrac{\phi_1}{1-\phi_2} T_i f_\alpha \le \dfrac{\phi_1}{1-\phi_2} f_{\alpha i}^{*}$. This result joined with Equation 6.16 implies $f_{\alpha i}^{*} \le \dfrac{\phi_1}{1-\phi_2} f_{\alpha i}^{*}$ which directly yields $(1-\phi_1-\phi_2) f_{\alpha,i}^{*} \le 0$.

Since $(1-\phi_1-\phi_2)>0$, $f^*_{\alpha,i} \leq 0$ and $f^*_i \leq \alpha$, i.e., arbitrarily poor local optima of class M_1 cannot exist.

Let us now consider M_2. Any $x^*_i \in M_2$ with value $f^*_{\alpha i}$ satisfies Equation 6.13 and Equation 6.17.

- If $\phi_1 \geq 0$, the analysis is identical to the case when $x^*_i \in M_1$ and the same conclusion is reached.

- If $\phi_1 \leq 0 \, and \, \phi_1+\phi_2 < 0$. Multiplying Equation 6.17 by $-\phi_1$ and adding $T_i^2 f_\alpha$ to both sides yields $T_i^2 f_\alpha - \phi_1 f^*_{\alpha i} \leq T_i^2 f_\alpha - \phi_1 T_i f_\alpha = \phi_2 f^*_{\alpha i}$ which implies $f^*_{\alpha i} \leq \frac{1}{\phi_1+\phi_2} T_i^2 f_\alpha$. If the two-step neighborhood average, $T_i^2 f_\alpha$, is nonnegative, $f^*_{\alpha,i} \leq 0$ and $f^*_i \leq \alpha$, $f^*_i \leq \alpha$, i.e., arbitrarily poor local optima of class M_2 cannot exist.

- If $\phi_1 \leq 0 \, and \, \phi_1+\phi_2 > 0$ it is easily shown that $f^*_{\alpha i} \geq \frac{1}{\phi_1+\phi_2} T_i^2 f_\alpha$. If the two-step neighborhood average, $T_i^2 f_\alpha$, is nonnegative, then $f^*_{\alpha,i} \geq 0$ and $f^*_i \geq \alpha$. A landscape, with respect to the set M_2, is *not favorable* for direct search methods.

We now summarize the above results for all one-step local minima, the x^*_i. Four cases exist:

1) If $\phi_1 \geq 0$, $\phi_1 + \phi_2 > 0$, and $T_i^2 f_\alpha \leq 0$ for *all* $x_i^* \in M_1 \cup M_2$, then $f_i^* \leq \alpha$, i.e., no arbitrary poor one-step local minima can exist.

2) If $\phi_1 \geq 0$, $\phi_1 + \phi_2 < 0$, and $T_i^2 f_\alpha \leq 0$ for *all* $x_i^* \in M_1 \cup M_2$, then every one-step local minimum has $f_i^* \equiv \alpha$, i.e., the global minimum is α.

3) If $\phi_1 \leq 0$, $\phi_1 + \phi_2 < 0$, and $T_i^2 f_\alpha \geq 0$ for *all* $x_i^* \in M_2$, then no arbitrarily poor one-step local minima, $x_i^* \in M_1 \cup M_2$, can exist.

4) If $\phi_1 \leq 0$, $\phi_1 + \phi_2 > 0$, and $T_i^2 f_\alpha \geq 0$ for *all* $x_i^* \in M_2$, then arbitrary poor local minima, $x_i^* \in M_2$, will exist (f_i^* *will* exceed α for one or more i) and the landscape is not favorable for a local search.

6.3. The Characteristic Landscape Equation for AR(p) Landscapes

Here we extend the development of the characteristic landscape equation for AR(2) landscapes to AR(p) landscapes.

Theorem 6.3: If the time series based on the random walk on the landscape is AR(p) process, then the landscape satisfies the equation

$$(T^p - \phi_1 T^{p-1} - \phi_2 T^{p-2} - \ldots - \phi_{p-1} T) f_\alpha = \phi_p f_\alpha \qquad (6.18)$$

Proof: The proof follows immediately by using the recurrence equation for AR(p) process

$$z_t = \phi_1 z_{t-1} + \phi_2 z_{t-2} + ... + \phi_p z_{t-p} + a_t$$

and applying the technique used in the proof of Theorem 6.1. □

Theorem 6.4: If the landscape satisfies the equation

$$(T^p - \phi_1 T^{p-1} - \phi_2 T^{p-2} - ... - \phi_{p-1} T) f_\alpha = \phi_p f_\alpha$$

then the time series based on a random walk on this landscape is consistent with AR(p) process.

Proof: The proof follows directly from the proof of Theorem 6.2. □

Next Chapter 7 contains some suggestion and ideas for further research.

Chapter 7

Suggestions for Further Research and Concluding Remarks

The accomplishments of this dissertation research, stated in Chapters 3 through 6, provide the basis for many different avenues of further research. Intriguing directions for future research related to Chapter 3 are:

(1) studying the properties of asymmetric neighborhoods which yield complex eigenvalues and eigenvectors. The use and meaning of such eigenvalues and eigenvectors has not been previously addressed.

(2) studying elementary landscapes where $\alpha \neq \mu$.

(3) developing methods to determine what neighborhood or neighborhoods would yield an elementary landscape for a stipulated solution space and objective function.

Chapter 4's research can be extended by deriving general guidelines on sufficient lengths of time series required to detect meaningful departures from AR(1) behavior. One extension of Chapter 5's work would be to develop a statistical method to estimate the lower (upper) bounds for local minima (maxima) associated with elementary landscapes. An extension of Chapter 6's efforts would be the investigation of the properties of AR(p) ($p>2$) landscapes using the characteristic landscape equation developed therein. Developing characteristic landscape equations for additional Box-Jenkins models in the MA, ARMA, and ARIMA classes and studying their possibly beneficial properties would also be a worthwhile endeavor.

Little research has been directed toward sampling methodologies for verifying whether a particular landscape is elementary. Stadler and Schnabl (1992) mention that small modification to the basic Traveling Salesman Problem (TSP) can cause significant differences in the performance of alternative solution techniques. No previous research has studied metrics to measure departures from a pristine elementary landscape. For example, we might start from an elementary landscape and change one or more values of the objective function vector. How we can measure how different the resulting landscape is from the "parent" elementary landscape. One approach might be to measure disparities in sample statistics yielded from a sample random walk on each of the pair of landscapes.

Another way to address the problem is to view the elementary landscape as simultaneous systems of equations (Laplacian equations) that have a specific solution. If we change some element of the parent elementary landscape the previous solution will no longer be globally valid. A sampling method might be developed to detect and measure how far the perturbed landscape is from a exact elementary state.

The research documented in this dissertation provides an initial foundation for achieving an understanding of the topology associated with the solution space of complex combinatorial optimization problems from the perspective of metaheuristic search methodology. Current understanding of such topology is largely based on empirical observation. It is my hope that future research will build upon this foundation and make inroads into scientific understanding why metaheuristic search methods work so well.

References

Angel, E., Zissimopoulos, V., *On the classification of NP-complete problems in terms of their correlation coefficient,* Discrete Applied Mathematics, (1999).

Barnes, J. W., Colletti, B., *Local Search Structure in the Symmetric Traveling Salesman Problem Under a General Class of Rearrangement Neighborhoods,* Applied Mathematics Letter, 14 (1), 105-108, (2001).

Barnes, J.W., S. Dokov, B. Dimova, and A. Solomon, *A Theory of Elementary Landscapes*, Applied Mathematics Letters, 16, (2003).

Berman, A., Plemmons, R. J., *Nonnegative Matrices in the Mathematical Sciences,* SIAM, Philadelphia, PA, (1994).

Biggas, N., *Algebraic Graph Theory,* Cambridge Univ. Press, New York, (1974).

Bollobas, B., *Graph Theory-An Introductory Course,* Springer Verlag, New York, (1979).

Box, G. E. P. and G. M. Jenkins, *Time Series Analysis*, Holden-Day, San Francisco, CA, (1976).

Carlton, F., Barnes, J. W., *Solving the Traveling Salesman Problem with Time Windows Using Tabu Search*, Institute of Industrial Engineering Transactions,28, 617-629, (1996).

Chunk, F. R. K., *Spectral Graph Theory*, Volume 92 of CBMS, American Mathematical Society, Providence, RI, (1997)

Codenotti, B. and L. Margara. *Local Properties of Some NP-Complete Problems*. TR-92-021, International Computer Science Institute, University of California at Berkeley (1992).

Colletti, Bruce W. *Group Theory and Metaheuristics.* Ph.D. Dissertation, The University of Texas at Austin, (1999).

Cvetkovic, D. M., Doob, M., Sachs, H., *Spectra of Graphs,* Academic, (1980).

Dimova, B., Barnes, J.W., Popova, E., *Arbitrary Elementary Landscapes & AR(1) Processes*, Applied Mathematics Letters, (to appear).

Glover, Fred and M. Laguna, *Tabu Search,* Kluwer Academic Publishers, Boston, MA, (1997).

Golub, G. H., Van Loan, C. F., *Matrix Computations*, Baltimore, MD, (1983).

Grover, Lov K., *Local Search and Local Structure of NP-Complete Problems,* Operations Research Letters, 12, 235-243, (1992).

Harary, F., *Graph Theory*, Addison-Wesley Publishing Company, (1972).

Hordijk, W., *A Measure of Landscapes,* Evolutionary Computation, 4, 335-360, (1996).

Hordijk, W., Stadler, P. F., *Amplitude Spectra of Fitness Landscapes,* Journal of Complex Systems, 1, 39-66, (1998).

Jenkins, G. M., Watts, D. G., *Spectral Analysis and Its Applications*, Holden-Day, San Francisco, CA, (1968).

Jensen, P. A., Bard, J. F., *Operations Research Models and Methods,* John Wiley & Sons, Inc., (2003).

Kemeny, J. G., J. L. Snell, *Finite Markov Chains*, D. Van Nostrand Co., Inc., Princeton, N.J., (1960).

Kulkarni, V. G., *Modeling and Analysis of Stochastic Systems*, Chapman & Hall, (1995).

Lancaster, P., Tismenetsky, M., *The theory of Matrices*, Academic Press, San Diego, CA, (1985).

Law, A. M., Kelton, W. D., *Simulation Modeling and Analysis,* McGraw-Hill, Inc., (1991).

Reeves, C. R, *Modern Heuristic Techniques for Combinatorial Problems*, Blackwell Scientific Press, Oxford (1993).

Reeves, C. R, *Landscapes, Operators and Heuristic Search,* Annals of Operations Research, 86, 473-490, (1999).

Ross, S., *Probability models*, Academic Press, (1997).

Ross, S., *Stochastic Processes*, John Wiley & Sons, (1983).

Sherrington, D., Kirkpatrick, S., *Solvable Model of a Spin Glass,* Phys. Rev. Letter, 35, 1792-1796, (1975).

Snell, K., *Finite Markov Chains,* D. Van Nostrand Company, Inc., (1960).

Solomon, A., Barnes, J. W., Dokov S. P., Acevedo, R., *Weakly Symmetric Graphs, Elementary landscapes and the TSP*, Applied Mathematics Letter, 16, 401-407, (2003).

Stadler, P.F., Schnabl, W., *The Landscape of the Traveling Salesman Problem,* Physics Letter A, 161, 337-344, (1992).

Stadler, P.F., *Landscapes and Their Correlation Functions*, Journal of Mathematical Chemistry, 20, (1996).

Stadler, P. F., *Random Walk and orthogonal functions associated with highly symmetric graphs,* Discrete Mathematics, (1993).

Stadler, P. F., Happel, R., *Random Field Models for Fitness Landscapes,* Journal of Mathematical Biology, 38, 435-478, (1999).

Stadler, P. F., Happel, R., *Correlation Structure of the Landscape of the Graph-Bipartitioning-Problem,* Journal of Physics A: Mathematical and General, 25, (1992).

Stadler, P. F., *Linear Operators on Correlated Landscapes,* J. Phys.I France 4: 681-696 (1994).

Stadler, P.F., R. Seitz, and G.P. Wagner, *Population Dependent Fourier Decomposition of Fitness Landscapes over Recombination Spaces: Evolvability of Complex Characters*, Bulletin of Mathematical Biology., 62, (2000).

Stewart, W. J., *Introduction to the Numerical Solution of Markov Chains*, Princeton University, Princeton, NJ, (1994).

Vesselin K. Vassilev, Terence C. Fogarty, Julian F. Miller, *Information Characteristics and the Structure of Landscape*, Evolutionary Computation, 8(1): 31-60 (2000).

Weinberger, E., *Correlated and Uncorrelated Fitness Landscapes and How to Tell the Difference*, Biological Cybernetics, 63, (1990).

Weinberger, E., *Fourier and Taylor Series on Fitness Landscapes,* Biological Cybernetics, Springer – Verlag, (1991).

Weinberger, E. D., Stadler, P.F., *Why Some Fitness Landscapes are Fractal,* Journal of Theoretical Biology, 163, 145-275, (1993).

Weinberger, E. D., *Local Properties of Kauffman's N-k Model: A Tunably Rugged Energy Landscape,* Physical Review A, 44, (1991).

Vita

Boryana Slavcheva Dimova was born in Sofia, Bulgaria on September 02, 1963, the daughter of Snezhana and Slavcho Slavchevi. A 1981 graduated of National High School of Mathematics and Science, Sofia, Bulgaria, she entered Sofia University in Sofia, Bulgaria. She received a degree of Bachelor of Mathematics and in July 1986 a degree Master of Mathematical Modeling from the Sofia University, Sofia, Bulgaria. She worked in the Institute for Industrial Engineering, Sofia, Bulgaria as a Programmer/Data Analyst, in the Foreign Language Institute, Sofia, Bulgaria as a Lecture of Mathematics, in the Institute of Technology, Sofia, Bulgaria as a Lecturer of Applied Mathematics, in the National High School of Mathematics and Science, Sofia, Bulgaria as a Lecturer of Mathematics, and in Fox Chase Cancer Center, Philadelphia as a Programmer/Data Analyst. In August 2001 she began Doctoral Studies in the Operations Research and Industrial Engineering Program at the University of Texas, Austin, Texas.

Permanent Address:

Kalimanci 33, bl.101, vh. B, et.4, apt. 23, Sofia, 1505, Bulgaria

This dissertation was typed by the author.